$\left(i.e., \; E = \sum_i p_{oi} \frac{y_{ri}}{p_{ri}}\right)$, and $Z_{\alpha/2}$ is the 100(1- α/2)% percentile of the standard normal distribution.

The 100(1-α)% confidence interval for an SMR is given by the interval (SMR_{LCL}, SMR_{UCL}). Thus, if the SMR_{LCL} of a 100(1-α)% confidence interval is greater than 1 (or 100%), then the SMR is statistically significantly different (greater) than 1 (or 100%) implying that the number of observed cause-specific deaths (e.g., lymphoid cancer deaths) in the cohort is more than the number of expected cause-specific deaths (e.g., lymphoid cancer deaths) in the general population with similar demographics as the cohort. On the contrary, if the SMR_{UCL} of a 100(1-α)% confidence interval is less than 1 (or 100%), then the SMR is statistically significantly different (less) than 1 (or 100%) implying that the number of observed cause-specific deaths (e.g., lymphoid cancer deaths) in the cohort is less than the number of expected cause-specific deaths (e.g., lymphoid cancer deaths) in the general population with similar demographics as the cohort.

The US lymphoid cancer mortality rates used for the calculations of the expected number of lymphoid cancer deaths are given in Tables 33-37.

A3.4 Calculating the Expected Number of Cause-Specific Deaths in a Cohort Assuming that the Death Rate in the Cohort Increases with Cumulative Exposure

The SMR is the ratio of observed to expected number of deaths in a cohort. The expected number of deaths is calculated assuming that the hazard rate is the background hazard rate of the reference population. However, if the background hazard rate is assumed to be affected by exposure to a carcinogen via a multiplicative function, then the expected number of deaths can be calculated assuming that the hazard rate is the product of the background hazard rate of the reference population multiplied by the exposure-response function that modifies the background rates. That is, the expected number of cause-specific deaths in a cohort can be calculated as:

$$\sum_i p_{oi} \times RR(d_i) \times \frac{y_{ri}}{p_{ri}}$$

where p_{oi} is the number of observed person-years in the i-th stratum of the study group, y_{ri} is the number of observed deaths in the i-th stratum of the reference population, p_{ri} is the number of person-years in the i-th stratum of the reference population, and $RR(d_i)$ is the exposure-response function (rate ratio function) evaluated at cumulative exposure d_i.

Using this expected number of cause-specific deaths in a cohort, an SMR* and bounds on the SMR* can be calculated as follows:

$$SMR^{*} = \frac{\sum_i y_{oi}}{\sum_i p_{oi} \times RR(d_i) \times \frac{y_{ri}}{p_{ri}}}$$

Similarly, the lower and upper limits of the 100(1-α)% confidence interval can be calculated as follows:

$$SMR^{*}_{LCL} = \frac{Obs}{E^{*}} \times \left(1 - \frac{1}{9 \times Obs} - \frac{Z_{\alpha/2}}{3 \times \sqrt{Obs}}\right)^{3}$$

and

$$SMR^{*}_{UCL} = \frac{(Obs + 1)}{E^{*}} \times \left(1 - \frac{1}{9 \times (Obs + 1)} + \frac{Z_{\alpha/2}}{3 \times \sqrt{Obs + 1}}\right)^{3}$$

where SMR^{*}_{LCL} is the 100(1-α/2)% lower confidence limit on the SMR*, SMR^{*}_{UCL} is the 100(1-α/2)% upper confidence limit on the SMR*, *Obs* is the number of observed cause-specific deaths (e.g., lymphoid cancer deaths) in the study $(i.e., \ Obs = \sum_i y_{oi})$, *E** is the expected cause-specific deaths (e.g., lymphoid cancer deaths) derived from the reference population background rates multiplied by the exposure response function *RR(d_i)* $\left(i.e., \ E^{*} = \sum_i p_{oi} \times RR(d_i) \times \frac{y_{ri}}{p_{ri}}\right)$, and $Z_{\alpha/2}$ is the 100(1- α/2)% percentile of the standard normal distribution.

A3.5 References

Breslow, N. E. and N. E. Day (1987). Statistical Methods in Cancer Research. Volume II – The Design and Analysis of Cohort Studies. International Agency for Research on Cancer, Lyon, France.

Checkoway, H., N. E. Pearce, and D. J. Crawford-Brown (1989). Research Methods in Epidemiology. Oxford University Press, New York, USA

Rothman, K. J. (1986). Modern Epidemiology. Little, Brown and Company, Boston USA / Toronto Canada

Table 33: Lymphoid Cancer Mortality Rates in the U.S. Population for Each Calendar Year (1930-1972), Each Race, Each Sex and Each Age Group (number of lymphoid cancer deaths per 100,000)

Age Group (Years)	1930	1940	1950	1960	1968	1969	1970	1971	1972
White Males									
< 1	0.571574	0.571574	0.571574	0.952897	0.664582	0.193834	0.250050	0.264904	0.436483
1-4	0.889715	0.889715	0.889715	0.905855	2.716523	2.469136	2.639159	2.639196	1.416049
5-9	0.896007	0.896007	0.896007	0.792474	3.181767	3.222868	3.486584	3.365958	3.053435
10-14	0.808974	0.808974	0.808974	0.764426	1.743532	2.089818	1.892907	1.777729	1.573083
15-19	1.173753	1.173753	1.173753	1.302018	2.187854	2.304943	2.062410	1.853147	1.868520
20-24	0.779566	0.779566	0.779566	1.226909	1.853888	1.437771	2.074683	1.564349	1.969677
25-34	1.246367	1.246367	1.246367	1.348092	1.948938	1.826095	1.642713	1.866738	1.436086
35-44	2.822822	2.822822	2.822822	3.369977	4.096598	4.063587	3.427241	3.219945	3.996754
45-54	6.291235	6.291235	6.291235	8.459325	10.379543	10.326954	10.435895	10.292100	9.491327
55-64	13.704865	13.704865	13.704865	18.845992	25.093104	24.651811	25.357608	27.116973	25.569775
65-74	18.092659	18.092659	18.092659	32.706133	53.237410	51.595092	51.896786	51.955307	51.216641
75-84	18.992015	18.992015	18.992015	38.781214	82.331839	88.898757	86.483903	88.585069	91.555937
85+	11.917858	11.917858	11.917858	37.471858	104.761905	101.686747	87.071343	105.399568	117.052632
Other Race Males									
< 1	0.493869	0.493869	0.493869	0.000000	0.342912	0.334609	0.950275	0.958681	1.354541
1-4	0.506669	0.506669	0.506669	0.510781	1.218451	1.163832	1.553219	0.925069	0.722674
5-9	0.875629	0.875629	0.875629	0.460755	1.440733	1.962067	1.107201	1.724138	1.617251
10-14	0.419074	0.419074	0.419074	0.374631	1.760325	1.713909	1.412963	0.949367	1.501877
15-19	0.639471	0.639471	0.639471	0.878770	2.205882	1.334380	1.415189	1.505376	1.782042
20-24	1.159879	1.159879	1.159879	0.798062	2.016607	1.771872	1.024119	1.309635	0.886525
25-34	1.371643	1.371643	1.371643	1.371711	1.282051	1.747997	1.386486	1.828030	1.277139
35-44	2.362183	2.362183	2.362183	3.357051	3.718674	3.658537	4.072298	4.099678	5.229794
45-54	5.984989	5.984989	5.984989	9.095071	11.770245	10.925926	12.172295	10.151380	12.971078
55-64	11.279807	11.279807	11.279807	17.047913	29.750000	31.365314	28.395850	31.578947	26.004728
65-74	11.984811	11.984811	11.984811	22.473431	45.908184	51.185771	46.782908	52.000000	43.314501
75-84	11.892728	11.892728	11.892728	23.349211	61.827957	62.765957	67.857013	57.692308	68.202765
85+	0.000000	0.000000	0.000000	15.943369	58.536585	52.272727	59.543142	80.851064	63.829787
White Females									
< 1	0.372830	0.372830	0.372830	0.466696	0.703416	0.752196	0.595918	0.419701	0.461215
1-4	0.589370	0.589370	0.589370	0.382623	2.033672	1.985371	1.976859	1.656868	1.449532
5-9	0.369624	0.369624	0.369624	0.240952	2.059308	2.331391	2.528940	2.320938	1.828012
10-14	0.231579	0.231579	0.231579	0.417692	1.185724	1.195589	1.110161	1.276644	1.255995
15-19	0.258359	0.258359	0.258359	0.242587	0.965624	0.882056	1.138742	1.116447	1.150775

Age Group (Years)	1930	1940	1950	1960	1968	1969	1970	1971	1972
20-24	0.521598	0.521598	0.521598	0.538865	0.859182	0.643897	0.830949	0.817682	0.823469
25-34	0.792567	0.792567	0.792567	0.695775	0.815707	0.811284	0.990505	0.730055	1.008598
35-44	1.656499	1.656499	1.656499	2.209093	2.610084	2.225193	2.125844	2.257623	2.227040
45-54	3.927054	3.927054	3.927054	5.317963	7.310358	6.770297	6.805298	6.449242	6.650224
55-64	9.581633	9.581633	9.581633	13.184796	16.236934	16.778907	16.683520	16.793724	15.473466
65-74	13.471141	13.471141	13.471141	21.389945	33.714562	34.345683	35.204790	33.589547	36.741455
75-84	13.544646	13.544646	13.544646	28.303572	54.802432	54.652880	56.864558	57.238122	56.749460
85+	11.466575	11.466575	11.466575	23.163091	57.645467	65.772669	57.425086	62.057522	59.322034
Other Race Females									
< 1	0.490851	0.490851	0.490851	0.649642	0.000000	0.343348	0.327084	0.659039	0.695476
1-4	0.255302	0.255302	0.255302	0.425917	0.788782	1.171171	1.564646	1.022305	0.545455
5-9	0.373279	0.373279	0.373279	0.153607	0.524246	0.721311	1.050270	1.136364	0.814664
10-14	0.000000	0.000000	0.000000	0.281193	1.222826	0.991408	0.837986	1.144310	0.629327
15-19	0.302773	0.302773	0.302773	0.122783	0.642055	1.078582	0.663027	0.921986	0.679348
20-24	0.572140	0.572140	0.572140	0.142154	1.020408	0.287632	0.898678	0.583333	0.960769
25-34	0.686160	0.686160	0.686160	0.906197	1.654997	1.175015	0.652594	0.694444	0.986842
35-44	1.574455	1.574455	1.574455	3.092078	2.105978	2.642276	2.321355	2.675585	2.514891
45-54	4.516905	4.516905	4.516905	7.099807	9.083333	9.046455	8.699902	8.268934	8.308157
55-64	7.848951	7.848951	7.848951	10.717328	20.000000	16.902944	18.750576	20.582121	16.276704
65-74	5.746153	5.746153	5.746153	12.368748	30.629139	27.597403	28.920872	31.981279	33.027523
75-84	4.880954	4.880954	4.880954	16.111612	37.500000	33.333333	32.715935	35.000000	34.437086
85+	0.000000	0.000000	0.000000	12.414341	29.508197	33.846154	22.881259	42.465753	36.842105

Table 34: Lymphoid Cancer Mortality Rates in the U.S. Population for Each Calendar Year (1973-1981), Each Race, Each Sex and Each Age Group (number of lymphoid cancer deaths per 100,000)

Age Group (Years)	1973	1974	1975	1976	1977	1978	1979	1980	1981
White Males									
< 1	0.908058	0.224475	0.528294	0.300067	0.500615	0.358533	0.273877	0.132507	0.132064
1-4	2.244898	1.937849	1.833031	1.491692	1.211771	1.370124	1.234337	0.999559	1.346066
5-9	3.192572	3.142184	2.786254	3.041926	2.701618	2.013605	2.703456	2.514574	2.153795
10-14	2.131166	2.046687	1.720841	1.787372	2.181993	1.920932	1.734473	1.758458	1.563759
15-19	1.934907	1.908439	1.957140	1.817788	1.691974	1.677743	1.720171	1.719677	1.542872
20-24	1.456249	1.256932	1.508621	1.205242	1.383173	1.537081	1.481645	1.646638	1.395948
25-34	1.559640	1.639344	1.467136	1.432200	1.456079	1.578878	1.322802	1.543315	1.499603

Age Group (Years)	1973	1974	1975	1976	1977	1978	1979	1980	1981
35-44	3.285860	3.206107	3.239279	2.932876	2.984485	3.414495	3.156437	3.505926	3.005275
45-54	9.415647	10.002913	9.567420	9.625196	9.086395	9.480337	9.692479	9.433185	9.489925
55-64	24.776732	24.812299	25.402042	24.272853	24.671202	24.745497	24.588897	25.549930	25.109082
65-74	52.533589	52.720450	50.549249	52.758868	52.749171	53.199113	54.677339	54.513390	52.882396
75-84	91.595563	91.298812	90.050167	92.269737	90.846216	96.881248	98.868072	98.827567	99.726331
85+	109.183673	109.126214	119.074074	116.333938	119.789842	125.252525	135.008104	135.478217	128.314866
Other Race Males									
< 1	0.000000	0.350064	0.000000	0.686344	0.000000	0.952922	0.604677	0.000000	0.000000
1-4	0.890472	1.334520	1.432408	1.648352	0.925926	0.915751	0.896057	0.867085	1.145101
5-9	1.717033	1.670146	1.742160	1.098901	2.105978	1.683502	1.346801	0.799939	1.551788
10-14	1.607916	1.411909	0.973828	1.039755	1.363918	1.322418	0.890019	1.453699	1.239236
15-19	1.851852	1.726343	1.179392	1.390568	1.014925	1.410106	1.567034	1.377656	1.363956
20-24	1.528014	1.383238	1.242236	1.187825	1.275691	1.709986	1.058901	1.480282	1.175116
25-34	1.333333	1.145475	1.243243	1.379663	1.699854	1.661283	1.179554	1.310302	1.284428
35-44	3.903201	2.773498	3.506098	3.048327	3.537906	3.778866	3.653586	3.462009	4.639626
45-54	9.490940	13.356164	10.365336	10.867734	10.067114	9.468439	11.367381	10.689003	10.210284
55-64	27.570093	29.633867	29.319955	30.363036	28.862661	25.991649	29.183673	29.668996	26.891935
65-74	56.880734	54.821429	53.739130	53.962901	54.545455	58.582677	50.844854	58.720972	54.042417
75-84	73.991031	76.855895	66.115702	74.806202	81.992337	76.226415	78.651685	85.585907	93.874677
85+	64.583333	76.000000	75.925926	60.000000	82.142857	108.620690	106.779661	80.643834	104.987699
White Females									
< 1	0.559929	0.396269	0.479311	0.555150	0.302594	0.455050	0.361702	0.210232	0.139542
1-4	1.087926	1.337486	1.087164	1.130952	1.031553	1.022044	0.964947	0.643648	0.888346
5-9	2.089711	1.931242	1.779013	1.525870	1.558551	1.671667	1.377491	1.181182	1.282891
10-14	1.010913	1.042753	0.977275	0.935829	1.054746	0.896104	0.828655	0.922761	1.031858
15-19	1.049838	0.888990	0.972081	0.705803	0.887341	0.700328	0.797176	0.818234	0.945110
20-24	0.683717	0.843359	0.774256	0.900794	0.672464	0.716642	0.628578	0.724198	0.705556
25-34	0.861660	0.811775	0.928295	0.739332	0.837019	0.936504	0.798198	0.855556	0.724416
35-44	2.267551	2.112676	2.106728	1.792044	1.865996	1.696495	1.630139	1.887533	1.727053
45-54	6.246017	6.551095	6.287809	6.452209	6.487905	6.471816	6.256618	6.115654	5.936539
55-64	16.013353	16.622439	15.990803	16.423433	16.627989	16.348638	16.209867	16.803601	17.030421
65-74	34.125587	34.821812	32.178287	34.755847	34.549814	35.034501	35.199592	37.603777	35.889455
75-84	58.124174	58.643892	57.581864	61.363079	61.298077	61.771617	63.731992	67.535625	68.589388
85+	67.239636	66.761364	67.724868	67.617450	76.367962	76.519130	75.692964	84.172570	83.353422
Other Race Females									
< 1	0.718184	0.000000	0.000000	0.000000	0.000000	0.654986	0.311744	0.000000	0.000000
1-4	0.898473	0.450045	1.364877	0.372439	0.753296	0.279851	0.547445	0.795146	0.583260
5-9	0.966851	0.629811	1.190476	0.968188	0.959561	0.886767	0.752394	0.407426	1.169315

Age Group (Years)	1973	1974	1975	1976	1977	1978	1979	1980	1981
10-14	0.623053	0.992556	0.802965	0.745805	0.693569	0.960307	0.774693	0.642377	0.757866
15-19	0.786885	0.571429	0.803461	0.422705	0.774732	0.587544	0.815376	0.864307	0.402981
20-24	0.538462	0.591716	0.283487	0.683060	0.654879	0.758534	0.612745	0.654753	0.634340
25-34	0.677083	0.935961	0.836431	0.924296	0.962343	0.558659	0.833018	1.034294	0.828562
35-44	2.156863	2.450032	1.977041	2.114428	2.238355	2.231356	2.103468	2.399917	2.864034
45-54	9.830007	6.540698	9.305655	6.770099	8.432056	6.662088	8.316430	8.035665	6.734315
55-64	18.818819	17.543860	19.038643	20.702403	19.516562	20.555074	18.891688	19.739761	18.660537
65-74	37.037037	34.240688	32.088520	34.087883	32.101911	32.885086	35.924617	32.425347	40.174421
75-84	31.761006	36.445783	44.067797	45.212766	48.041775	45.641026	47.727273	57.289609	57.167055
85+	46.250000	54.117647	41.935484	43.877551	45.192308	50.000000	63.157895	65.743449	70.517392

Table 35: Lymphoid Cancer Mortality Rates in the U.S. Population for Each Calendar Year (1982-1990), Each Race, Each Sex and Each Age Group (number of lymphoid cancer deaths per 100,000)

Age Group (Years)	1982	1983	1984	1985	1986	1987	1988	1989	1990
White Males									
< 1	0.000000	0.462407	0.000000	0.192266	0.064567	0.512302	0.000000	0.244261	0.118477
1-4	0.897367	1.310122	0.781290	0.830986	0.877404	0.739505	0.737235	0.663349	0.708275
5-9	2.366171	1.846937	1.510829	1.428039	1.366221	1.467699	1.225459	1.297239	0.913484
10-14	1.583212	1.360994	1.426616	1.285190	1.274476	1.210121	1.201909	1.428199	1.352777
15-19	1.796605	1.780555	1.689925	1.682906	1.512290	1.333880	1.353366	1.212178	1.409300
20-24	1.343823	1.284539	1.270779	1.324499	1.419361	1.497749	1.274751	1.514134	1.248516
25-34	1.527609	1.570647	1.584635	1.706365	2.154965	1.607166	1.992268	1.977337	2.268786
35-44	3.607424	3.210907	3.607591	3.900018	3.907493	3.733309	3.744332	4.073447	3.925666
45-54	10.320582	9.492029	9.475140	9.981628	10.353269	10.305775	10.121232	10.454357	11.342008
55-64	25.740401	25.933995	26.359149	27.642635	26.093181	28.162326	28.577168	29.628210	29.421239
65-74	55.446249	58.683266	58.006916	60.547081	63.379973	61.768858	60.894609	63.835855	64.680548
75-84	102.512985	103.269530	102.903810	113.797884	111.957418	110.325657	117.539257	121.572182	124.689270
85+	141.091466	154.657919	146.182157	158.545624	152.478016	146.762825	171.258407	163.709977	185.700410
Other Race Males									
< 1	0.282407	0.000000	0.560626	0.544009	0.265887	0.513383	0.243094	0.231537	0.000000
1-4	0.950552	0.843139	0.898864	0.815968	0.584038	0.359246	0.352241	0.545662	0.529965
5-9	1.544365	1.263091	1.035059	1.065461	1.635687	1.002256	0.802618	0.847424	0.838924
10-14	1.101152	1.094825	1.341328	1.465289	1.305275	0.991744	0.674730	1.075256	0.990555
15-19	1.544260	1.214203	1.108428	0.701977	0.978176	1.531826	1.121842	1.232062	0.892218

Age Group (Years)	1982	1983	1984	1985	1986	1987	1988	1989	1990
20-24	0.848498	1.603323	1.108261	1.322919	1.200467	0.919044	1.446631	1.389804	1.442548
25-34	1.840239	1.941467	1.637358	1.906600	1.752430	1.457848	1.865610	2.782049	2.290311
35-44	3.630473	3.495188	4.120332	4.426983	4.713920	4.554605	4.972986	4.699949	5.240313
45-54	12.753297	11.795082	11.153652	10.804774	11.090469	11.424834	12.745138	13.021074	13.059052
55-64	27.441584	33.281437	30.656579	29.982650	30.277039	26.602320	29.171684	30.098894	33.984171
65-74	57.237298	55.381074	50.838187	61.469040	67.722773	64.142203	60.374990	60.402824	65.684984
75-84	99.028610	108.712639	94.311838	97.257155	112.593187	106.228728	99.871509	110.026091	109.071026
85+	110.976140	120.734757	82.336687	113.366296	106.579982	137.074874	121.273370	148.091471	159.703198
White Females									
< 1	0.412871	0.418804	0.207705	0.338393	0.204025	0.337325	0.397082	0.450230	0.062415
1-4	0.740887	0.943464	0.464971	0.714428	0.693092	0.601971	0.653006	0.419260	0.451249
5-9	1.294763	0.911457	0.835611	0.988693	0.757493	0.627520	0.559821	0.641137	0.623382
10-14	0.811883	0.631763	0.881446	0.834117	0.803605	0.716906	0.557631	0.640258	0.556603
15-19	0.816159	0.870140	0.723414	0.626600	0.838982	0.794999	0.644126	0.647127	0.788964
20-24	0.873275	0.679190	0.641055	0.778479	0.804127	0.708784	0.656806	0.791296	0.786603
25-34	0.743563	0.696736	0.814677	0.906247	0.940198	0.770082	0.829128	0.869329	0.884170
35-44	1.741456	1.859996	2.115381	1.992830	1.956782	1.717332	2.159311	1.856792	1.787279
45-54	6.734416	6.563147	6.457907	6.609959	6.253106	6.042936	6.355324	6.076045	6.084263
55-64	16.917034	17.085084	17.960658	18.684330	17.474939	17.735989	17.586514	18.798277	17.622023
65-74	37.596194	39.177268	39.824889	39.607408	41.121751	40.965889	41.342613	43.020215	43.082987
75-84	69.543091	70.552506	72.529403	71.315776	76.337351	76.845877	77.916555	80.989763	81.092049
85+	92.412534	89.912880	93.843998	94.727554	100.448726	104.084539	103.516519	109.816269	114.634887
Other Race Females									
< 1	0.292722	0.000000	0.868817	0.563369	0.553598	0.000000	0.252484	0.239977	0.468898
1-4	0.726035	0.546679	0.611366	0.454753	0.298587	0.515052	1.010791	0.699719	0.476427
5-9	0.548698	1.087145	0.198370	0.640049	0.804902	0.421807	0.645421	0.520951	0.458591
10-14	0.812410	0.622286	0.437587	0.752269	0.382603	0.509268	0.377932	0.490451	0.477840
15-19	0.580762	0.764674	0.593717	0.298791	0.471507	0.640464	0.461812	0.519634	0.748110
20-24	0.853074	0.561540	0.501356	0.221421	0.554927	0.671071	0.564213	0.510058	0.851649
25-34	0.731149	0.674739	0.950363	1.008959	0.926506	0.903771	1.071554	0.710502	0.963634
35-44	2.213313	2.192893	2.291606	2.543862	2.321505	2.242482	2.132750	2.326151	2.652870
45-54	7.298407	7.121108	7.312326	6.550464	8.025120	7.634042	7.331957	7.589449	8.253123
55-64	18.533248	17.381368	20.156957	19.876547	18.758072	18.216235	19.695708	19.588978	19.595873
65-74	37.355813	38.276541	36.088017	38.533843	40.391660	39.156632	40.894103	41.773392	41.612207
75-84	59.725264	61.003109	58.979590	72.662063	61.616938	61.855941	67.427820	70.322620	71.910686
85+	64.834220	66.926697	64.149876	77.144586	79.929917	83.506794	81.033922	81.645237	83.769867

Table 36: Lymphoid Cancer Mortality Rates in the U.S. Population for Each Calendar Year (1991-1999), Each Race, Each Sex and Each Age Group (number of lymphoid cancer deaths per 100,000)

Age Group (Years)	1991	1992	1993	1994	1995	1996	1997	1998	1999
White Males									
< 1	0.120549	0.304542	0.309342	0.250062	0.125911	0.126229	0.381286	0.313145	0.261647
1-4	0.598010	0.634873	0.641730	0.483114	0.597917	0.525628	0.322071	0.389179	0.520896
5-9	1.077332	1.046375	0.842215	0.869082	1.071523	0.627185	0.728541	0.635617	0.535847
10-14	1.069727	0.922609	1.018617	0.953443	0.855020	0.884591	0.804178	0.847763	0.589373
15-19	1.394160	1.411226	1.281312	1.131257	1.049657	1.046720	0.934061	1.187142	0.880738
20-24	1.486628	1.485252	1.049435	1.532901	1.098601	1.291260	1.508268	1.552742	1.398208
25-34	2.153514	2.230164	2.090814	2.252798	2.244475	2.011220	2.201578	1.773869	1.305571
35-44	4.716193	4.434700	4.386889	4.381832	4.635446	4.322717	3.891075	3.694620	2.936410
45-54	11.299132	10.765887	10.498471	11.240728	10.956518	10.384872	10.941259	10.085568	9.264970
55-64	28.990578	28.964490	28.869688	30.789233	30.267561	29.977605	29.599598	28.278056	27.768360
65-74	65.820142	67.437957	67.622686	70.574494	70.831434	69.983251	72.455585	71.013446	69.063573
75-84	123.244041	128.192453	129.169255	130.541394	132.139030	135.097298	134.542905	135.014407	136.039499
85+	184.620012	182.774888	186.482519	202.084388	203.049861	205.679170	195.813850	199.761637	200.496795
Other Race Males									
< 1	0.000000	0.000000	0.231198	0.000000	0.490283	0.492542	0.242734	0.476757	0.000000
1-4	0.251040	0.180786	0.291989	0.172394	0.286071	0.287824	0.233362	0.352567	0.176170
5-9	0.706327	0.689215	0.565082	0.492402	0.520381	0.819514	0.572628	0.430521	0.256131
10-14	0.775427	0.641820	0.568414	0.759836	1.047504	0.733418	0.767420	0.561479	0.813209
15-19	1.191880	1.185346	0.500675	0.864956	1.198790	0.553187	0.731660	0.662851	1.070727
20-24	1.124612	1.642354	1.785301	1.508855	0.972847	1.313934	2.015238	0.645289	0.993891
25-34	2.237519	2.484545	2.407845	2.206208	2.567098	2.425574	2.111731	1.761624	1.717844
35-44	5.264830	5.221627	4.846035	4.669117	5.130747	5.026924	5.259584	4.383872	3.907748
45-54	12.192547	12.871079	12.740362	12.099461	12.981341	12.574332	13.039173	11.972081	9.760551
55-64	31.597492	34.051901	28.743845	34.058142	31.510938	32.051830	30.667501	30.433409	31.292855
65-74	67.516141	61.893730	69.133246	62.181494	62.604246	67.819297	64.586214	62.510594	61.446247
75-84	118.346204	108.465272	111.503892	101.134128	110.952607	117.171986	116.895856	108.432653	108.149986
85+	131.534134	140.571056	164.607271	156.009507	161.524956	154.217709	152.287127	162.763360	161.416252
White Females									
< 1	0.189610	0.128216	0.260841	0.394373	0.198615	0.463996	0.600393	0.328510	0.206611
1-4	0.544654	0.484663	0.362290	0.393668	0.231834	0.268299	0.322384	0.375495	0.411102
5-9	0.617083	0.712038	0.651712	0.505619	0.510744	0.422820	0.559046	0.412139	0.282375
10-14	0.420396	0.650159	0.510683	0.558181	0.525734	0.507201	0.530655	0.539522	0.375783
15-19	0.791386	0.689823	0.563043	0.653104	0.495588	0.564889	0.605686	0.474534	0.521361

Age Group (Years)	1991	1992	1993	1994	1995	1996	1997	1998	1999
20-24	0.719853	0.647753	0.577305	0.783432	0.732804	0.840555	0.913694	0.930414	0.701500
25-34	0.928258	0.984040	0.944766	1.037638	0.882957	1.072279	0.822517	0.832823	0.824799
35-44	1.920846	1.937426	1.865423	2.084310	2.097702	1.968226	1.983071	1.727557	1.672751
45-54	6.500862	5.997125	5.912764	6.459897	6.114375	6.139397	5.639134	5.577498	5.202266
55-64	19.178724	18.330817	19.220898	19.593339	19.239323	19.268723	19.531043	17.763069	17.363737
65-74	44.670651	45.063962	46.706389	46.334466	47.634353	46.662600	47.170072	45.873513	46.282577
75-84	85.652607	85.539274	87.768235	88.536784	89.289949	90.527655	89.550870	91.065418	91.226321
85+	118.035157	115.502420	120.620701	117.264248	125.040442	121.648591	124.871721	121.364315	122.155611
Other Race Females									
< 1	0.234086	0.000000	0.000000	0.000000	0.254598	0.254855	0.504694	0.000000	1.249619
1-4	0.193747	0.434289	0.180589	0.415097	0.472506	0.356208	0.300468	0.120879	0.181199
5-9	0.502308	0.109141	0.688359	0.355915	0.489020	0.376693	0.364674	0.178399	0.221088
10-14	0.340783	0.658581	0.265457	0.260343	0.718685	0.604677	0.148552	0.193467	0.420867
15-19	0.760147	0.290665	0.629617	0.667091	0.589240	0.516753	0.551219	0.243356	0.478619
20-24	0.552215	0.701958	0.744932	0.369962	0.529128	0.641656	0.371187	0.574389	0.811758
25-34	1.250760	1.161703	1.074879	0.969668	1.282122	1.191926	1.034714	1.221072	0.860489
35-44	2.631571	2.695297	2.201742	2.072282	2.737377	2.480527	2.904835	2.831665	2.114252
45-54	7.433460	7.524094	7.964662	7.841874	7.423539	6.577967	6.862564	6.910658	6.250333
55-64	20.877164	19.463921	21.271408	20.568934	23.617713	21.535597	20.943180	21.726642	21.037674
65-74	46.704315	41.136051	43.407193	39.603040	41.951707	46.011816	43.479905	44.474852	41.977259
75-84	81.049219	72.227947	77.173631	76.716888	75.573071	76.119672	72.954561	78.245435	76.115208
85+	87.337153	99.305842	94.501598	94.680398	94.904241	99.516750	98.701031	99.677092	95.995562

Table 37: Lymphoid Cancer Mortality Rates in the U.S. Population for Each Calendar Year (2000-2008), Each Race, Each Sex and Each Age Group (number of lymphoid cancer deaths per 100,000)

Age Group (Years)	2000	2001	2002	2003	2004	2005	2006	2007	2008
White Males									
< 1	0.524806	0.250750	0.381423	0.126342	0.125603	0.063462	0.378854	0.433816	0.375811
1-4	0.390715	0.311593	0.340849	0.547846	0.383588	0.428761	0.414535	0.207105	0.460199
5-9	0.647961	0.536133	0.544783	0.809098	0.738830	0.586288	0.440868	0.721561	0.485417
10-14	0.836564	0.644528	0.792704	0.683952	0.508571	0.705677	0.615860	0.597909	0.405742
15-19	1.143733	1.118192	1.005208	0.941732	1.015803	0.933706	0.867502	0.827787	0.838181
20-24	1.424321	1.262936	1.335348	1.160621	1.051160	1.247020	1.314343	1.043871	1.270049
25-34	1.207456	1.325997	1.292035	1.232081	1.287954	1.026088	1.180857	1.123533	1.249620

Age Group (Years)	2000	2001	2002	2003	2004	2005	2006	2007	2008
35-44	2.951331	2.947883	2.787913	2.719071	2.445056	2.470472	2.151277	2.365903	2.161794
45-54	8.736368	8.658735	8.160044	7.522465	7.274624	6.838794	6.861847	6.613099	6.164806
55-64	26.024599	25.768249	24.602045	24.337611	22.290379	21.443948	20.815903	20.218269	20.093016
65-74	68.210725	66.846157	66.754466	63.724138	59.058038	59.772839	55.443301	55.225882	52.210701
75-84	137.861646	131.603614	132.026187	129.571266	125.750437	126.843740	126.655258	125.431566	123.714919
85+	202.953378	206.959834	212.138265	213.290538	201.174047	212.220517	195.502713	202.949122	202.726728
Other Race Males									
< 1	0.235491	0.000000	0.448970	0.000000	0.000000	0.211882	0.207428	0.000000	0.389636
1-4	0.232676	0.174487	0.114159	0.281887	0.388513	0.436998	0.324330	0.529700	0.359809
5-9	0.426663	0.433151	0.350934	0.177529	0.536648	0.669715	0.307361	0.432344	0.255016
10-14	0.352086	0.844244	0.697316	0.803100	0.437740	0.359507	0.481909	0.444312	0.486827
15-19	0.920683	1.076046	0.792248	0.602980	0.459569	0.604006	0.779758	0.720078	0.890076
20-24	1.679528	1.056120	0.877657	1.167735	1.357733	1.165263	1.232959	1.051449	0.744980
25-34	1.363152	1.404313	1.538684	1.551104	1.403061	1.602819	1.098655	1.126761	1.266334
35-44	2.835120	3.817562	3.392236	3.049851	2.553021	2.602693	3.074193	3.089058	2.116457
45-54	10.717689	9.866223	8.851983	9.939288	9.058168	9.391368	8.899028	8.540407	7.925244
55-64	26.363186	29.985785	26.175855	23.212888	23.481933	23.096876	24.894886	21.742272	21.917414
65-74	61.467682	61.255497	57.822519	52.268589	57.715894	54.302768	52.212361	49.404447	51.758535
75-84	102.947245	104.276589	99.069233	95.457067	100.239504	96.713415	94.921776	97.159675	93.011377
85+	145.308316	142.557723	134.973258	143.433958	145.190271	126.514193	152.502927	143.278205	131.946501
White Females									
< 1	0.483682	0.131239	0.332853	0.596126	0.263276	0.199731	0.198550	0.324862	0.327583
1-4	0.376789	0.310412	0.392293	0.388978	0.217928	0.199665	0.334287	0.317396	0.216318
5-9	0.425186	0.446824	0.547368	0.446350	0.436685	0.356507	0.299872	0.379088	0.375590
10-14	0.486294	0.377656	0.561295	0.397890	0.411565	0.441312	0.381939	0.540134	0.375560
15-19	0.492428	0.502412	0.435949	0.420339	0.629975	0.422781	0.479903	0.488373	0.438460
20-24	0.606969	0.729405	0.791141	0.676381	0.607536	0.555826	0.530911	0.682503	0.390786
25-34	0.751260	0.854954	0.782482	0.621166	0.630221	0.725255	0.731735	0.641508	0.582598
35-44	1.522875	1.588986	1.609632	1.453520	1.243847	1.286495	1.359781	1.251519	1.204327
45-54	5.326357	4.737304	4.630905	4.389539	4.295574	3.898529	3.933733	3.694953	3.534546
55-64	17.389128	16.335271	15.009996	13.676430	13.322191	13.352400	12.130725	11.797667	11.197640
65-74	44.010466	41.752191	40.585987	37.403030	36.937724	35.289786	35.434227	33.258375	31.591145
75-84	90.119912	87.396791	84.699781	84.711257	82.164651	81.038234	78.777329	78.024018	75.235482
85+	128.513697	128.834098	129.776449	128.647982	124.750168	125.342160	126.731086	123.320293	121.223154
Other Race Females									
< 1	0.244260	0.000000	0.464279	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1-4	0.359362	0.179663	0.176290	0.232051	0.114423	0.000000	0.000000	0.164215	0.053145
5-9	0.309062	0.402679	0.271573	0.228525	0.459707	0.000000	0.135214	0.266130	0.261604

Age Group (Years)	2000	2001	2002	2003	2004	2005	2006	2007	2008
10-14	0.227928	0.174845	0.254859	0.499492	0.206140	0.289557	0.373864	0.083534	0.377093
15-19	0.520827	0.465908	0.824630	0.536728	0.260194	0.208961	0.283326	0.236140	0.231250
20-24	0.838657	0.702065	0.770675	0.398600	0.393036	0.650290	0.687329	0.466475	0.581676
25-34	1.000629	1.272210	1.020700	0.869944	0.899656	0.752461	0.696625	0.664100	0.611427
35-44	2.317793	2.049276	1.899200	1.862371	1.737403	2.008196	1.872617	1.809375	1.348465
45-54	6.319216	6.213190	6.929462	5.666120	5.479445	5.300950	5.361658	5.400012	4.546107
55-64	17.592975	18.765077	17.788091	14.672254	15.503902	15.881942	14.640494	14.890397	13.472998
65-74	40.580024	41.223164	41.278055	41.797987	36.900825	36.086683	34.291068	34.010516	31.508649
75-84	74.119505	74.499069	70.453876	77.651645	71.641475	61.796102	62.880913	66.641937	62.963260
85+	115.616309	97.336673	86.333420	98.078476	99.450371	89.589566	92.445974	88.253258	86.059963

Appendix 4 Hypothetical Example of Appearance of Supra-Linearity in the Absence of Truly Low-Dose Data

USEPA acknowledges that "the actual exposure-response relationship at low exposure levels is unknown" (pp. 4-61 and 4-74 of USEPA 2016). The inability to observe sublinearity in the NIOSH cohort might be explained by the lack of dose-response data at low air concentrations (e.g., a few ppb) that would allow total internal exposures (endogenous + exogenous) to remain in/near the normal endogenous range (e.g., see Figures 3 and 6). Where available dose-response data are predominated by exposures above the area in the dose-response expected to be sublinear (i.e., within/near/below the normal endogenous range in the present case), if the doses are sufficiently high to be in the area of the dose-response where disproportionally increased risk occurs, then the dose-response observed based on the data available might appear supra-linear overall. As a hypothetical example, Figure 14 below is similar to Figure 4-2 of USEPA (2016) for lymphoid cancer.

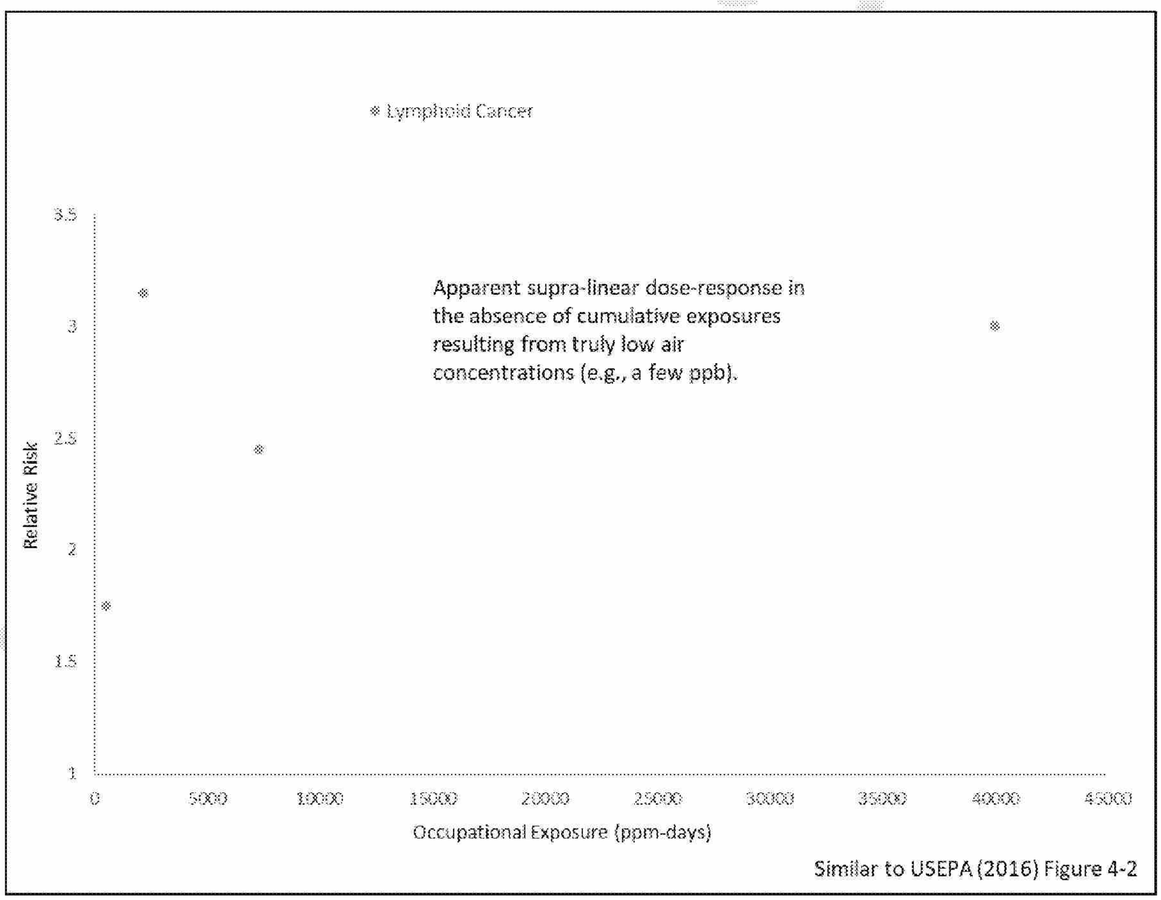


Figure 14: Seemingly Supra-linear Dose-Response for Lymphoid Cancer

The dose-response as presented (not based on the individual data or additional exposure groups) may appear overall supra-linear in nature, as noted by USEPA (2016). However, examination of the dose axis reveals that there are no truly low-dose data to characterize the shape of the dose-response at low exposures (e.g., a few ppb), especially within/near/below the endogenous range where USEPA (2016) expects sublinearity. Hypothetical dose-response data in the range of endogenous exposures and below were used to produce Figures 15 and 16 (see below).

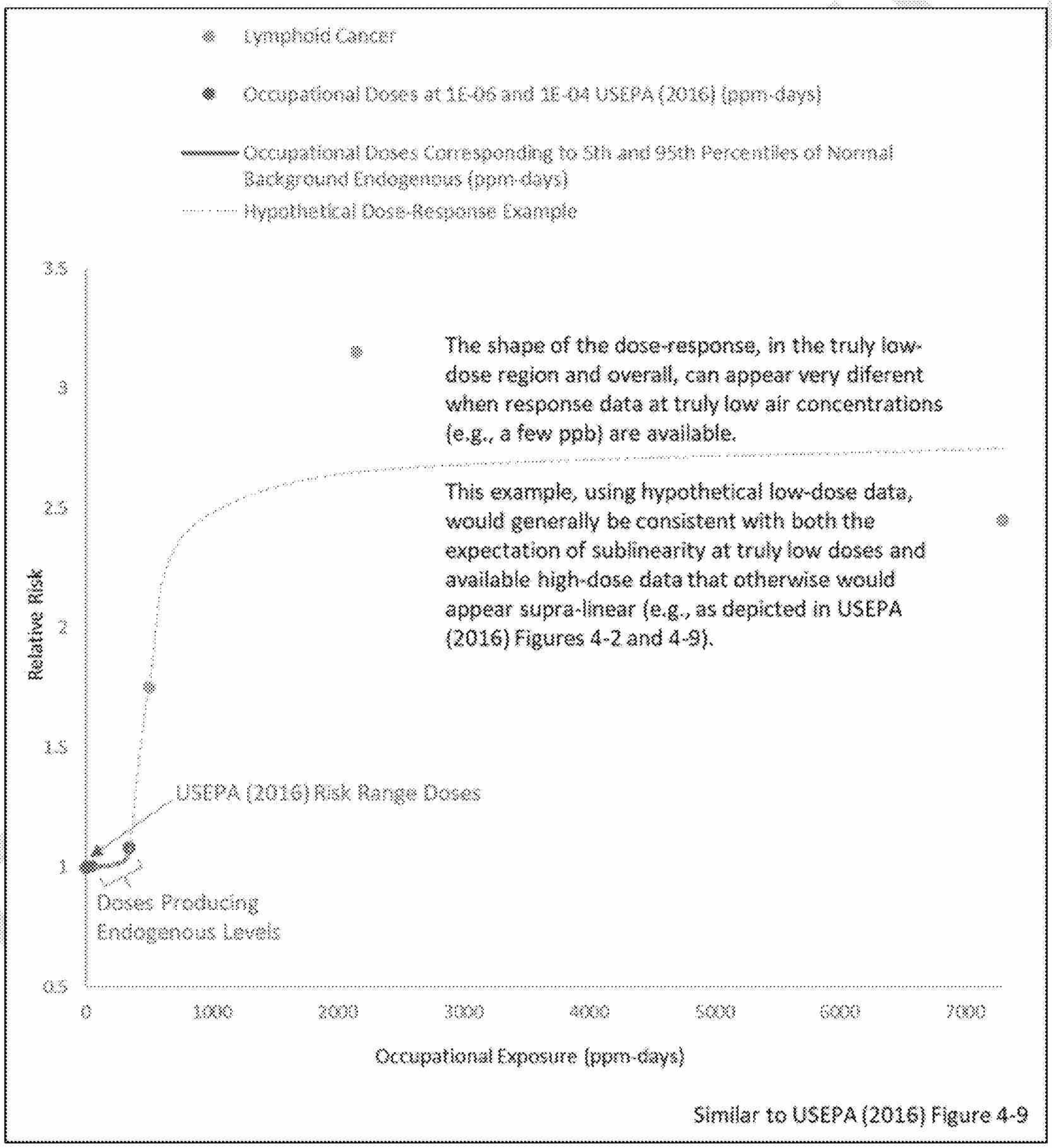


Figure 15. Hypothetical Sublinear Dose-Response at Truly Low Doses Plotted with Available High-Dose Data for Lymphoid Cancer

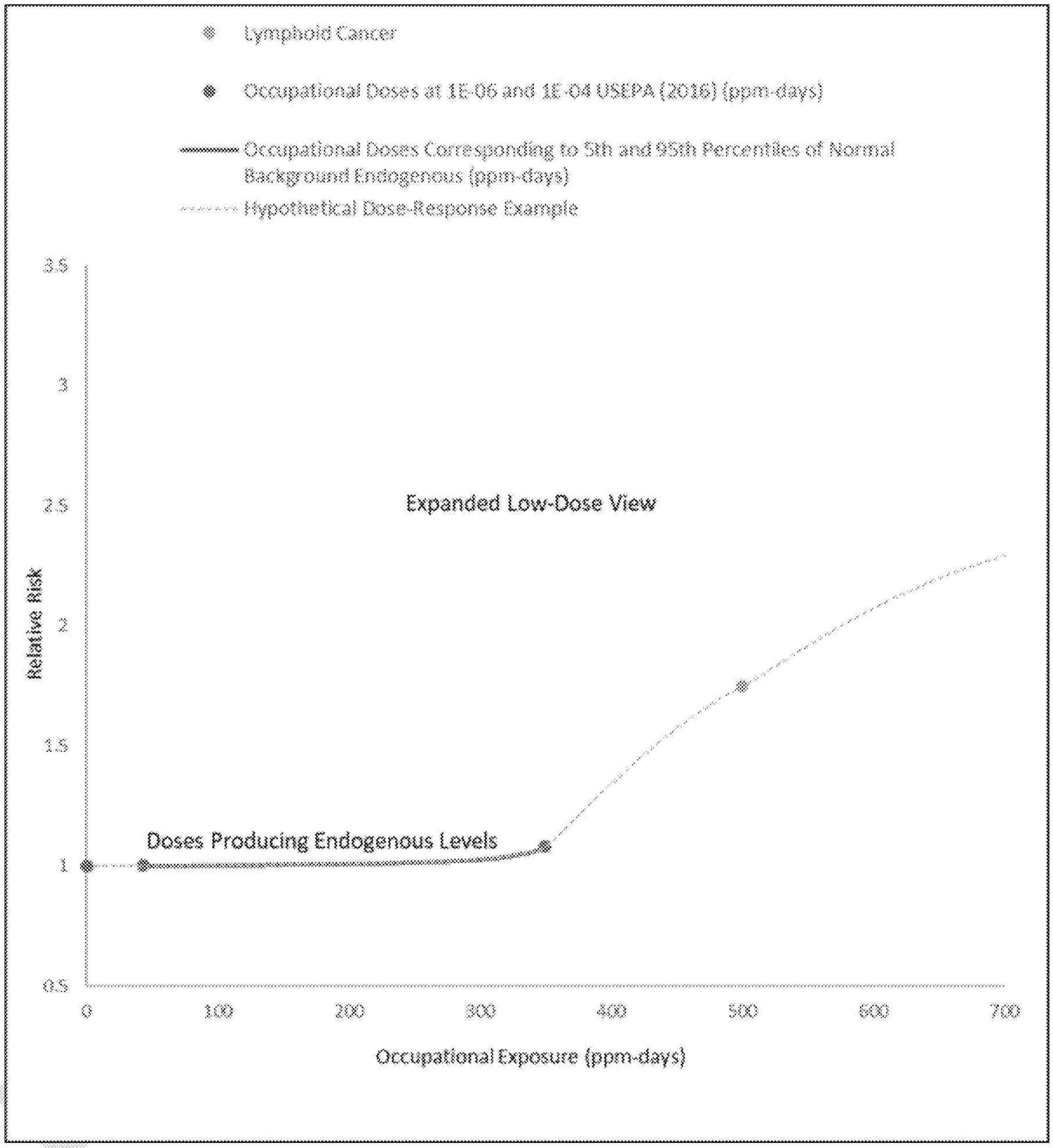


Figure 16: Hypothetical Sublinear Dose-Response at Truly Low Doses Plotted with Available High-Dose Data for Lymphoid Cancer – Expanded Low-Dose View

The availability of adequate, truly low-dose data in this hypothetical example reveals the existence of sublinearity in the overall dose-response at doses corresponding to the endogenous range (and significantly lower doses corresponding to 1E-06 to 1E-04 excess risk based on USEPA 2016). However, simple removal of these truly low-dose data results in a graph depicting a seemingly supra-linear dose-response (Figure 17) with a steep low-dose slope down to a relative risk of 1 at 0 dose (similar to Figure 4-9 in USEPA 2016). At the same time, it should be realized that use of a different (e.g., higher) number of cumulative exposure intervals provides a different visual impression (e.g., see Figure 6S of Valdez-Flores et al. 2013).

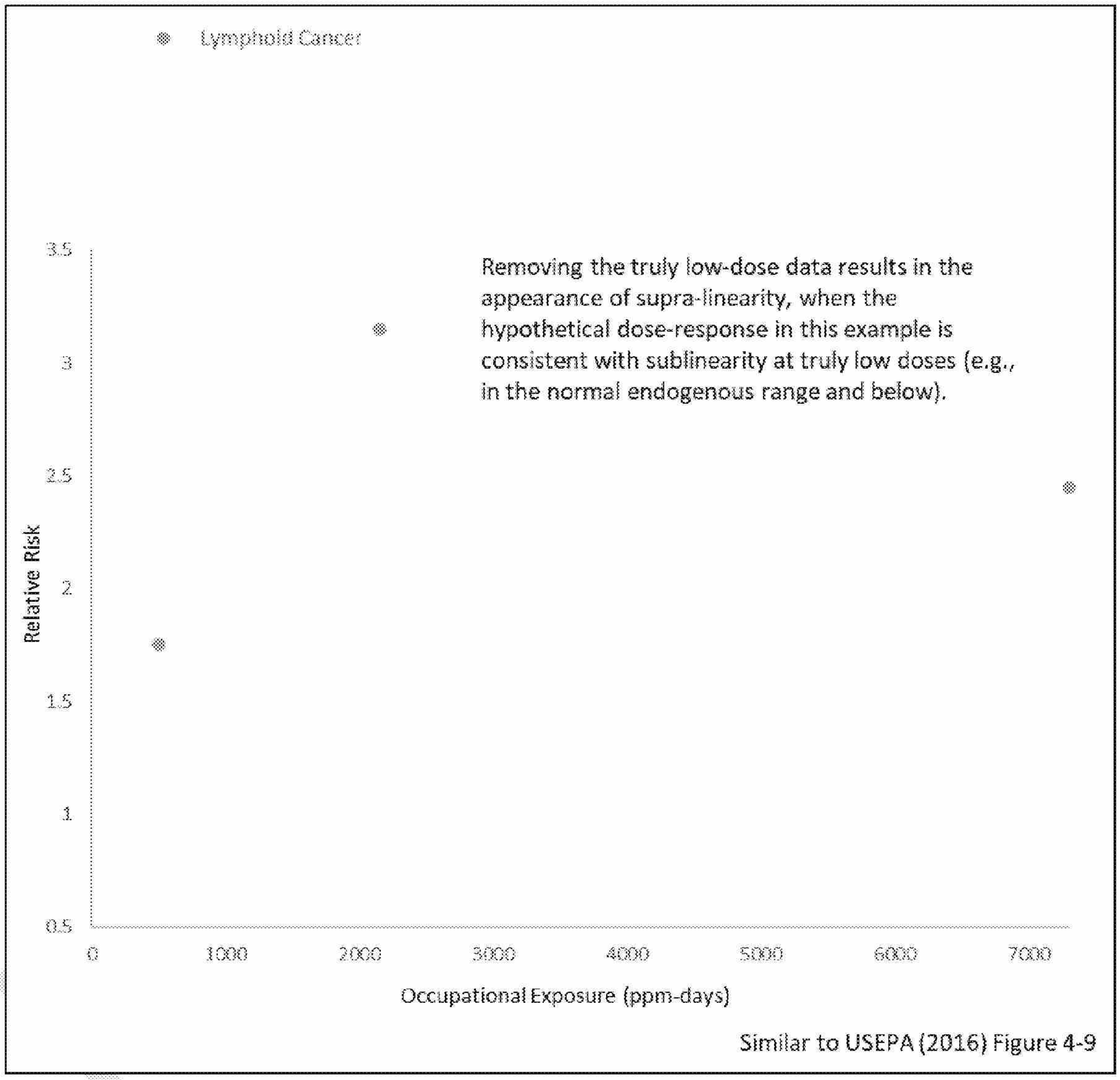


Figure 17: Seemingly Supra-linear Dose-Response from Removal of Hypothetical Low-Dose Data for Lymphoid Cancer

Figure 18 also depicts the possibility of a downward shift in the apparent dose-response curve in the absence of truly low-dose data, where the dose range for the apparent supra-linear curve on the left could be similar to that in Figure 14.

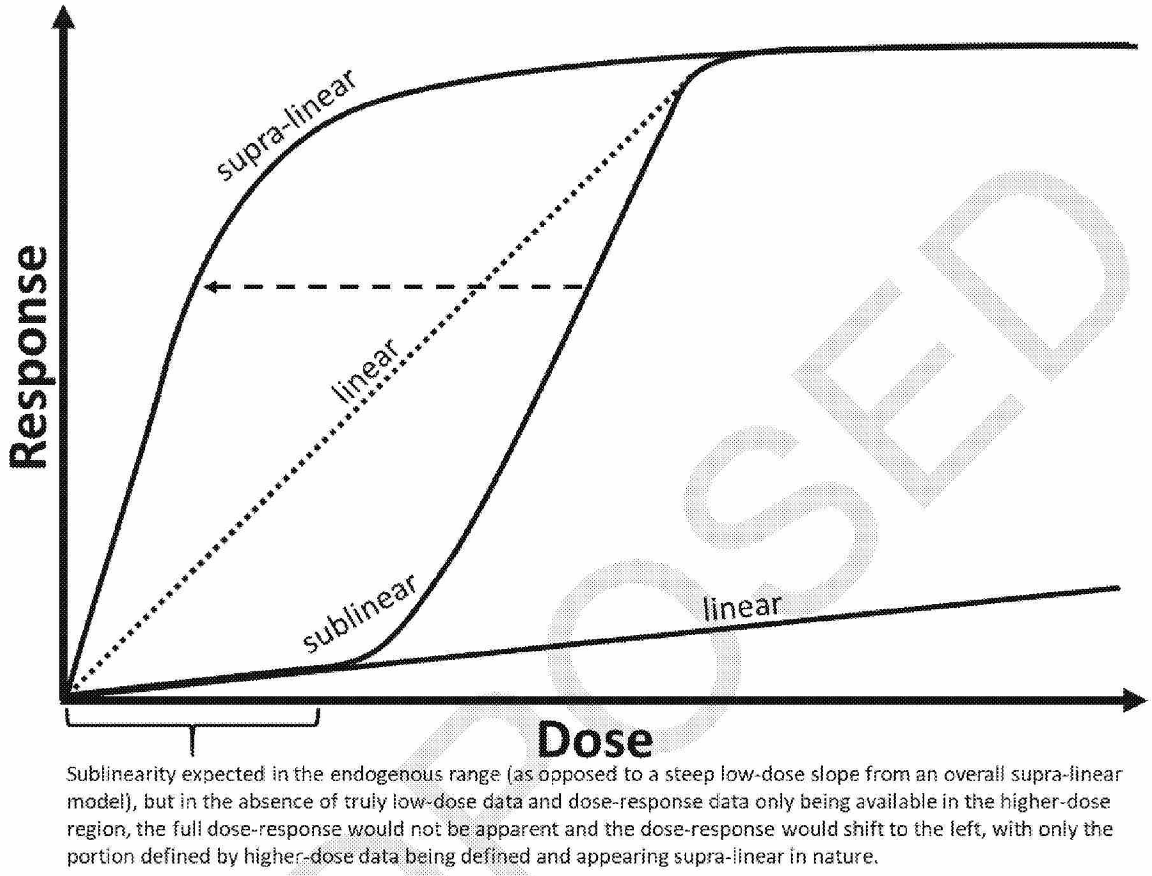


Figure 18: Seemingly Supra-linear Dose-Response from Removal of Hypothetical Low-Dose Data

These examples simply demonstrate the hypothetical possibility of the appearance of an overall supra-linear dose-response, despite an underlying true dose-response that is sublinear at truly low doses, *when available data are at relatively high doses* above the sublinear portion of the curve and into the steep slope portion wherein high response per unit dose is induced.

To help put the high occupational EtO exposures into perspective, the environmental-to-occupational level corresponding to 1E-04 excess risk based on USEPA (2016) is 0.777 ppm-days (i.e., 0.00001 ppm (environmental at 1E-04 excess risk) × 70 years × 365 days/year × 20 m^3/10 m^3 × 365 days/240 days = 0.777 ppm-days occupational). By comparison, the midpoint of the lowest exposure group for lymphoid cell lineage cancer in Steenland et al. (2004) is almost 800 times lower (≈600 ppm-days; 15-year exposure lag) and had an odds ratio of only 0.90. Even today the OSHA PEL (1 ppm) is 222 times the air concentration corresponding to the 95th percentile of the normal endogenous background range (4.5 ppb; Table 4 of Kirman and Hays

2017), and around 294,000-417,000 times higher than central tendency environmental levels (i.e., background and environmental exposure means ≈0.0044-0.0062 µg/m^3 (0.0024-0.0034 ppb) per USEPA 2016).

The TCEQ has not evaluated the hypothetical above further as it is somewhat beyond the scope of this DSD.

Appendix 5 Corrected p-Values and Akaike Information Criterion (AIC) for the Two-Piece Spline Model and Other Models

A5.1 Lymphoid Cancer:

A5.1.1 Corrected p-value example for the log-linear spline model with knot at 1,600 ppm-days

The likelihood ratio test is used to test whether a fitted model significantly improves the fit of the data by estimating parameters instead of just assuming a baseline (null) model for the data. The likelihood ratio test is evaluated by comparing the likelihood of the model with the estimated parameters and the likelihood of the null model. If the likelihood of the model with the estimated parameters is equal to the likelihood of the null model, then the natural logarithm of the ratio of these likelihoods multiplied by two follow a Chi-Square distribution with as many degrees of freedom as the number of parameters estimated for the fitted model. Thus, if the fit of the baseline (null) model and the model with estimated parameters are not different,

$$Chi - Square(k) = \chi_k^2 = -2\ln\left(\frac{likelihood\ for\ null\ model}{likelihood\ for\ fitted\ model}\right)$$

This can also be written as follows,

$$\chi_k^2 = -2LogL(null\ model) + 2LogL(fitted\ model)$$

Here k is the number of degrees of freedom (k is the number of parameters that were estimated in excess of the parameters estimated for the null model or nested model).

For the log-linear spline model with knot at 1,600 ppm-days for lymphoid cancer (Table D-33 on p. D-46 of USEPA 2016), the χ_k^2 value was equal to 5.2722 (463.912-458.640) and k was set to 2. This resulted in a p-value of 0.0716. That is, the fitted model was assumed to have two parameters; namely, the slope below the knot and the slope above the knot. The results are from a Statistical Analysis System (SAS) output for the model specified. The two-piece log-linear model specified included a knot. This knot was determined so that the likelihood of the spline model was maximized. That is, the knot is another parameter that was searched for outside SAS. Because the estimation of the knot was done outside SAS, the SAS program did not count the knot as a parameter and, consequently, the Chi-Square test SAS reported does not reflect the fact that the knot was also estimated. The correct Chi-Square that accounts for the fact that the knot was estimated outside SAS should then be 5.2722, but k (the degrees of freedom) should be three. This corrected calculation would result in a p-value of 0.1529. That is, the corrected p-value indicates that the likelihood of the log-linear spline model with knot at 1,600 ppm × days is not different from the likelihood of the null model at the 5% significance level. In plain words, there is not enough evidence indicating that the fitted two-piece log-linear spline model explains the variability in the data any better than the null model. The same is true for the linear two-piece spline model with a "knot" at 1,600 ppm × days selected by USEPA (p-value of 0.14).

A5.1.2 Corrected AIC value example for the log-linear spline model with knot at 1,600 ppm-days

The AIC is equal to 2*k* - 2*LogL* where *k* is the number of parameters estimated for the model and LogL is the logarithm of the likelihood. Table D-33 in USEPA (2016) lists the -2LogL as 458.640 and the AIC as 462.640. That is:

$$462.640 = 2k + 458.640$$

The AIC and –2LogL implies that k equals 2. That is, the spline model was assumed to have estimated two parameters; namely, the slope below the knot and the slope above the knot. The results in Table D-33 (p. D-46 of USEPA 2016) consist of SAS output for the two-piece log-linear spline model specified. The model specified included a knot. This knot was pre-assigned (i.e., previously estimated using a separate optimization procedure outside the SAS run), so the likelihood of the model was maximized only *conditional on* the estimated knot-value used for that calculation. Consequently, the knot must be treated as an additional parameter that was estimated outside SAS. Because the estimation of the knot was done outside SAS, the SAS run performed by USEPA did not count the knot as a model parameter and, consequently, the resulting AIC value it obtained does not reflect that the knot was in fact estimated. USEPA could have requested SAS to account properly for the extra degree of freedom properly associated with its estimated knot value, but USEPA evidently elected not to make this request of SAS.

The correct AIC, which accounts for the fact that the knot was estimated outside SAS, should instead be:

$$AIC = 464.640 = 2 \times 3 + 458.640$$

Correct AIC values and p-values for all models in Table 4-6 of USEPA (2016) are summarized in the corrected USEPA Table 4-6 below, which is Table 38 of this DSD (i.e., the p-values and AIC values have been corrected to reflect the degree of freedom for the knot in the two-piece spline models and to reflect the likelihood difference between SAS procedures used for linear and log-linear models).

Table 38: Corrected USEPA Table 4-6 - Models Considered for Modeling the EtO Exposure-Response Data for Lymphoid Cancer Mortality in Both Sexes in the NIOSH Cohort for the Derivation of Unit Risk Estimates

Model [a]	p-value [b]	AIC [c]	USEPA Comments
Two-piece spline models			
Linear spline model with knot at 1,600 ppm × days	0.14	*464.5*	**SELECTED.** Adequate statistical and visual fit, including local fit to low-exposure range; linear model; AIC within two units of lowest AIC of models considered.
Linear spline model with knot at 100 ppm × days	0.11	*463.8*	Good overall statistical fit and lowest AIC of two-piece spline models, but poor local fit to the low-exposure region, with no cases below the knot.
Log-linear spline model with knot at 1,600 ppm × days	0.15	464.6	Linear model preferred to log-linear (see text above).
Log-linear spline model with knot at 100 ppm × days	0.11	463.8	Good overall statistical fit and tied for lowest AICc of two-piece spline models, but poor local fit to the low-exposure region, with no cases below the knot.
Linear (ERR) models (RR = 1 + β × exposure)			
Linear model	0.13	*463.6*	Not statistically significant overall fit and poor visual fit.
Linear model with log cumulative exposure	0.02	*460.6*	Good overall statistical fit, but poor local fit to the low-exposure region.
Linear model with square-root transformation of cumulative exposure	0.053	*462.2*	Borderline statistical fit, but poor local fit to the low-exposure region.
Log-linear (Cox regression) models ($RR = e^{\beta \times exposure}$)			
Log-linear model (standard Cox regression model)	0.22	464.4	Not statistically significant overall fit and poor visual fit.
Log-linear model with log cumulative exposure	0.02	460.4	Good overall statistical fit; lowest AIC[c] of models considered; low-exposure slope becomes increasingly steep as exposures decrease, and large unit risk estimates can result; preference given to the two-piece spline models because they have a better ability to provide a good local fit to the low-exposure range.
Log-linear model with square-root transformation of cumulative exposure	0.08	462.8	Not statistically significant overall fit and poor visual fit.

[a] All with cumulative exposure as the exposure variable, except where noted, and with a 15-yr lag.
[b] p-values from likelihood ratio test, except for linear regression of categorical results, where Wald p-values are reported. $p < 0.05$ considered "good" statistical fit; $0.05 < p < 0.10$ considered "adequate" statistical fit if significant exposure-response relationships have already been established with similar models.
[c] AICs for linear models are directly comparable and AICs for log-linear models are directly comparable. However, for the lymphoid cancer data, SAS proc NLP (where NLP = nonlinear programming) consistently yielded −2LLs and AICs about 0.4 units lower than proc PHREG for the same models, including the null model, presumably for computational processing reasons, and proc NLP was used for the linear RR models. Thus, AICs for linear models

are equivalent to AICs about 0.4 units higher for log-linear models. No AIC was calculated for the linear regression of categorical results. *In order to make the AICs comparable for different models, the AICs for the linear models have been increased by 0.4 to reflect the discrepancy in the -2LogL values reported by the SAS proc NLP and by SAS PHREG (as italicized in this table).*

Table 38 shows that *neither the linear two-piece spline model with a "knot" at 1,600 ppm × days selected by USEPA (2016) nor the standard Cox regression model fit the data statistically significantly better than the null model* (zero slope). Additionally, the AIC values are very similar. *However, as use of a supra-linear model (i.e., the steep lower-dose slope) is not scientifically justified for low-dose extrapolation (see Section 3.4.1.4.1), the two-piece spline models are not considered for adoption*; nor are other models that have an inherently supra-linear dose-response over the exposure range (i.e., log-linear or linear models with log cumulative exposure or with square-root transformation of cumulative exposure). As for the linear model, it neither fits the data statistically better than the null model (at the 5% significance level) nor is consistent with USEPA's/TCEQ's expectation of sublinearity in the endogenous range, while the standard Cox regression model is consistent. Lastly, *no superior model fit is readily apparent visually based on accurate depictions of model fit to the actual underlying data* (Appendix 6). Thus, based on these and other considerations discussed in this DSD, *the TCEQ selects the standard Cox regression model for lymphoid cancer mortality*.

A5.2 Breast Cancer Incidence

A5.2.1 Corrected AIC example for the linear spline model with knot at 5,750 ppm-days

Similar to Table 38 above for lymphoid cancer, correct AIC values and p-values for all breast cancer incidence models in Table 4-14 of USEPA (2016) are summarized in the corrected USEPA Table 4-14 below, which is Table 39 in this DSD (i.e., the p-values and AIC values have been corrected to reflect the degree of freedom for the knot in the two-piece spline models and to reflect the likelihood difference between SAS procedures used for linear and log-linear models).

Table 39: Corrected USEPA Table 4-14 - Models Considered for Modeling the EtO Exposure-Response Data for Breast Cancer Incidence in Females in the Subcohort with Interviews from the NIOSH and Health Incidence Study Cohort for the Derivation of Unit Risk Estimates

Model [a]	p-value [d]	AIC [b]	USEPA Comments
Two-piece spline models			
Two-piece linear spline model (knot at 5,750 ppm × days)	*0.0367*	1,956.360 [c]	**SELECTED**. Good overall statistical fit and good visual fit, including local fit to low-exposure range; linear model; AIC within two units of lowest AIC of models considered.
Two-piece log-linear spline model (knot at 5,800 ppm × days)	*0.0384*	*1,956.485*	Good overall statistical fit and good visual fit, including local fit to low-exposure range; preference given to the two-piece linear spline model primarily because it has the advantageous property of linearity, but it also has a marginally better statistical fit (lower AIC).
Linear (ERR) models (RR = 1 + β × exposure)			
Linear model with square-root transformation of cumulative exposure	*0.0038*	1,952.501	Good overall statistical fit and lowest AIC; low-exposure slope becomes increasingly steep as exposures decrease, and large unit risk estimates can result; preference given to the two-piece spline models because they have a better ability to provide a good local fit to the low-exposure range.
Linear model with untransformed cumulative exposure	*0.0114*	1,954.526	Good overall statistical fit but poorer local fit to low-exposure range than the two-piece spline models; higher AIC than selected model.
Log-linear (Cox regression) models (RR = $e^{\beta \times \text{exposure}}$)			
Log-linear model with square-root transformation of exposure	*0.0049*	1,953.028	Good overall statistical fit; low-exposure slope becomes increasingly steep as exposures decrease, and large unit risk estimates can result; preference given to the two-piece spline models because they have a better ability to provide a good local fit to the low-exposure range.
Log-linear model with (natural) log cumulative exposure	*0.0302*	1,956.176	Good overall statistical fit but poor local fit to low-exposure range; low-exposure slope becomes increasingly steep as exposures decrease, and large unit risk estimates can result; higher AIC than selected model.
Log-linear model (standard Cox regression)	0.0404	1,956.675	Good overall statistical fit but poor local fit to low-exposure range (too shallow); AIC exceeds that of selected model by >2.
Linear regression of categorical results			
Linear regression of categorical results, excluding the highest exposure quintile	---	--- [c]	Not statistically significant, as one might expect because the approach, which is based on categorical data, has low statistical power; preference given to models that treated exposure as a continuous variable and that also provided reasonable representations of the low-exposure region.

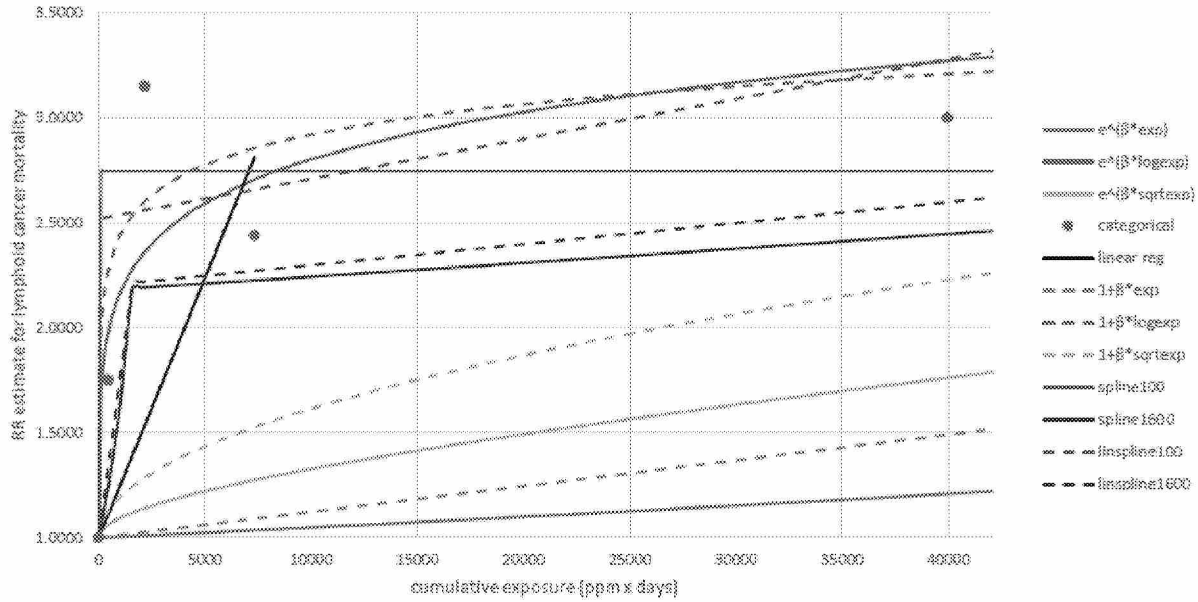
[a] All with cumulative exposure as the exposure variable, except where noted, and with a 15-yr lag, and all with exposure as a continuous variable except for the linear regression of categorical results.
[b] AIC = 2p-2LL, where *p* = number of parameters and LL = ln(likelihood), assuming two exposure parameters for the two-piece spline models.
[c] Not calculated.

[d] *p-values were calculated from EPA's Table D-2.*
[e] *AIC values for the two-piece spline models were adjusted to reflect the degree of freedom for the knot.*

Table 39 shows that *both the linear two-piece spline model with a "knot" at 5,750 ppm × days selected by USEPA (2016) and the standard Cox regression model selected by the TCEQ fit the data statistically significantly better than the null model* (zero slope). Additionally, the AIC values are very similar. *However, as use of a supra-linear model (i.e., the steep lower-dose component) is not scientifically justified for low-dose extrapolation (see Section 3.4.1.4.1), the two-piece spline models are not considered for adoption*; nor are other models that have an inherently supra-linear dose-response over the exposure range (i.e., log-linear or linear models with log cumulative exposure or with square-root transformation of cumulative exposure). While the linear model is not consistent with USEPA's/TCEQ's expectation of sublinearity in the endogenous range, the standard Cox regression model is consistent. Thus, based on these and other considerations discussed in this DSD, *the TCEQ will consider standard Cox regression modeling results for breast cancer incidence.*

Appendix 6 Visual Fit to the Underlying NIOSH Data

Visual fit to the data was used by USEPA (2016) as a criterion for model selection. However, no appropriate visual comparison of model fit to the lymphoid cancer mortality data can be made based on Figure 4-3 (p. 4-21 of USEPA 2016) since the data shown are not even the data to which the models were fit. As such, USEPA Figure 4-3 (shown below as Figure 19 of this DSD) misrepresents model fit.



e^(β*exp): RR = $e^{(\beta \times \text{exposure})}$; e^(β*logexp): RR = $e^{(\beta \times \ln(\text{exposure}))}$; e^(β*sqrtexp): RR = $e^{(\beta \times \text{sqrt}(\text{exposure}))}$; categorical: RR = $e^{(\beta \times \text{exposure})}$ with categorical exposures, plotted at the mean cumulative exposure; linear reg: weighted linear regression of categorical results, excluding highest exposure group (see text); 1 + β*exp: RR = 1 + β × exposure; 1 + β*logexp: RR = 1 + β × ln(exposure); 1 + β*sqrtexp: RR = 1 + β × sqrt(exposure); spline100(1,600): Two-piece log-linear spline model with knot at 100 (1,600) ppm × days (see text); linspline100(1,600): Two-piece linear spline model with knot at 100 (1,600) ppm × days (see text). (Note that, with the exception of the categorical results and the linear regression of the categorical results, the different models have different implicitly estimated baseline risks; thus, they are not strictly comparable to each other in terms of RR values, i.e., along the *y*-axis. They are, however, comparable in terms of general shape.)

Source: Steenland reanalyses for males and females combined; see Appendix D (except for linear regression of categorical results, which was done by EPA).

Figure 4-3. Exposure-response models for lymphoid cancer mortality vs. occupational cumulative exposure (with 15-year lag).

Figure 19: USEPA (2016) Figure 4-3

More specifically, the actual data underlying the model fits are the individual data, not the less refined categorical data shown in USEPA Figure 4-3. *Thus, because the model fits shown in USEPA (2016) Figure 4-3 are those to the individual data (and not the categorical data depicted), the figure does not actually indicate model fit to the modelled data at all.*

Categorical rate ratios (RRs) should not be used for visually comparing models fit to individual data, particularly when appropriate statistical model fit criteria are available. More specifically, estimated nonparametric RRs are calculated with respect to an underlying background hazard rate that is also estimated nonparametrically. The RRs of parametric models fit to the individual data are defined with respect to an underlying background hazard rate estimated by the model. However, the underlying background hazard rates estimated by the nonparametric RRs and the parametric model are generally different. A better comparison of models fit to the observed data is to use the predictiveness of the model; that is, the capability of the model to estimate the observed number of deaths with a certain degree of confidence (see Appendix 3). Moreover, visual interpretation of the consistency of categorical RRs with the shape/slope of a modelled dose-response can change as the number of exposure categories changes. For example, Figures 1-3 of Valdez-Flores and Sielken (2013) demonstrate, among other things, how the dose-response (i.e., dose-RR) slope for breast cancer mortality in the NIOSH cohort appears very steep when compared to only four exposure categories but seems more shallow when additional categories are added (i.e., up to 20 and 61 categorical RRs). In the present case, the overall dose-response appears ill represented by only a few categorical RRs, whether for breast cancer (see Figures 1-3 of Valdez-Flores and Sielken 2013) or lymphoid cancer (see below and supplementary material for Valdez-Flores and Sielken 2013).

The visual presentation of only a few exposure categories can blind the data user to the variability in the underlying dose-response data, and by corollary, preclude an appropriate visual assessment/comparison of model fit to the actual individual data. For example, in looking at all lymphoid cancer death RRs for the NIOSH cohort in the Figures 20-22 below (e.g., as opposed to a few categorical RRs represented by the red dots), objective examination of the model fits to the underlying data reveals no readily apparent superior fit by any particular model. What is most readily apparent is the loss of visualized information that results from only using the five grouped RRs (represented by the red dots) as in Figure 4-3 of USEPA (2016). The nonparametric rate ratios for individual cases (categorical) represented by the black circles in Figure 22 below form no discernable pattern that appears most consistent with any specific model (i.e., visual fit cannot be used to readily identify a model fit most representative of the actual data). In fact, other dose-responses could be added that would appear equally plausible and/or consistent with these high-dose occupational data.

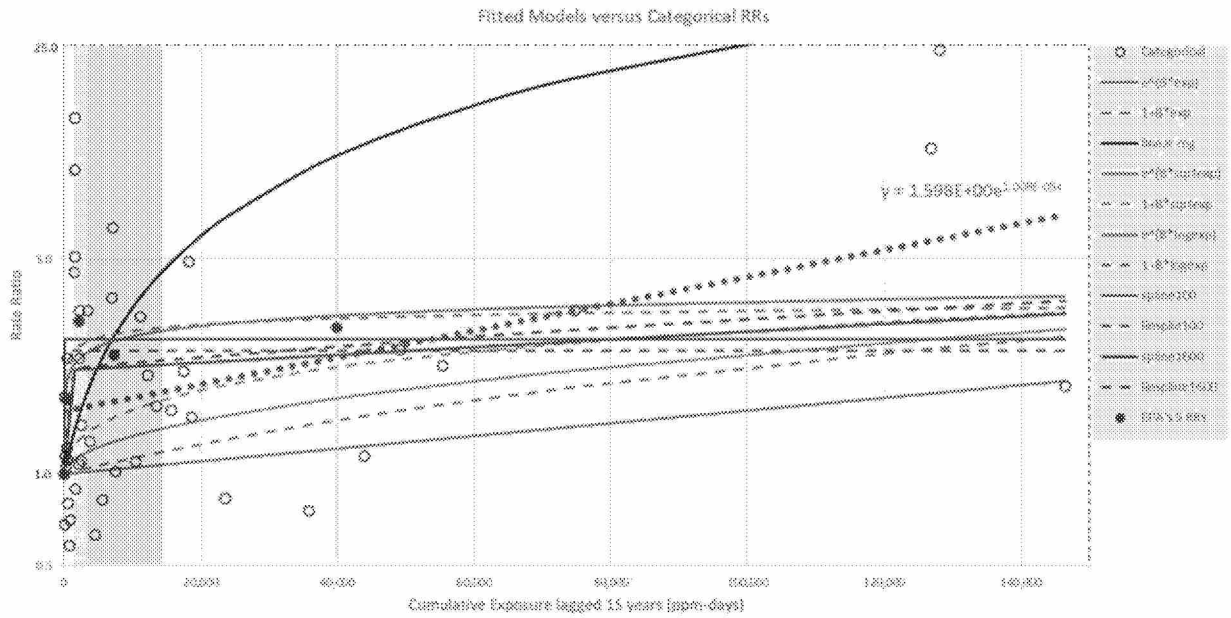


Figure 20: Lymphoid Cancer Death Categorical Rate Ratios (RRs) and Various Fitted Models for 15-Year Lagged Occupational Doses ≤150,000 ppm × days (NIOSH cohort)

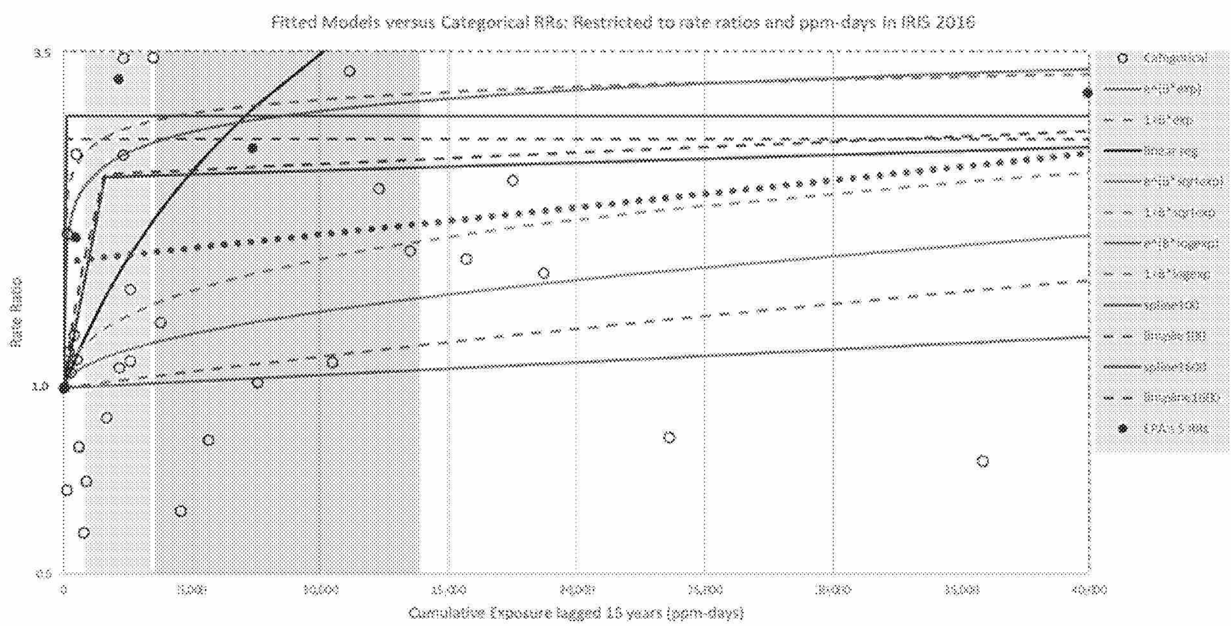


Figure 21: Lymphoid Cancer Death Categorical RRs and Various Fitted Models for 15-Year Lagged Occupational Doses ≤40,000 ppm × days (NIOSH cohort)

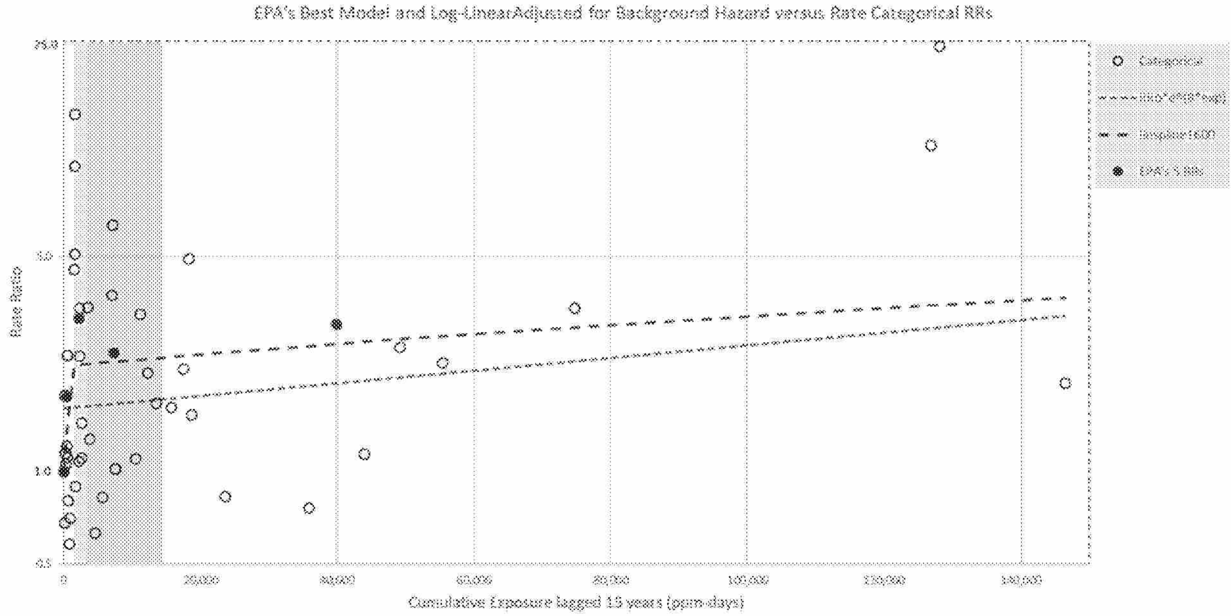


Figure 22: Lymphoid Cancer Death Categorical RRs and the Cox Proportional Hazards and Two-Piece Spline ("knot" at 1,600 ppm × days) Fitted Models for 15-Year Lagged Occupational Doses ≤150,000 ppm × days (NIOSH cohort)

The TCEQ contends that no defensible model fit conclusions may be drawn based on these more elucidating, transparent, and accurate depictions of model fit other than no superior model fit is readily apparent (e.g., see the three figures above and Figure 3 of Valdez-Flores and Sielken 2013). Appropriate statistical model fit criteria are evaluated elsewhere in this DSD (Section 3.4.1.3 and Appendix 5).

[Note: In these graphs, the dotted light blue line approximates the correct visual representation of the log-linear model (standard proportional hazards model) fit to the full NIOSH dataset after adjusting for the difference in baseline risks between the rate ratios and the log-linear model, thereby addressing USEPA's following footnote to Figure 4-3 (p. 4-21 of USEPA 2016) concerning the visual incomparability of model fit to the data, "Note that, with the exception of the categorical results and the linear regression of the categorical results, the different models have different implicitly estimated baseline risks; thus, they are not strictly comparable to each other in terms of RR values, i.e., along the y-axis." The model "RRo* e^(B*exp)" is an approximation of the log-linear model (e^(B*exp)) adjusted through multiplying by the ratio of the underlying baseline hazard rate of the model to the underlying baseline hazard rate the nonparametric estimates.]